PCT/EP 97/02443

09/180798

The Patent Office

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

PRIORITY DOCUMENT

Signed

Dated 26-3-1997

Patents Form 1/77

Patents Act 1977
(Rules 16)

The Patent Office

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

14 MAY 1996

1. Your reference — 137-1103/SG

2. Patent application number *(The Patent Office will fill in this part)* — **9610044.1**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

SANDOZ LTD.
35 Lichtstrasse
CH-4002 Basel
Switzerland

Patents ADP number *(if you know it)* 00703207001

If the applicant is a corporate body, give the country/state of its incorporation — Switzerland

4. Titel of the invention — IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

5. Name of your agent *(if you have one)* — B. A. Yorke & Co.

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Coomb House
7 St. John's Road
Isleworth, Middlesex TW7 6NH

Patents ADP number *(if you know it)* 1800001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from a earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer "Yes" if:*

a) any applicant named in part 3 is not an inventor , or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

Patents Form 1/77

Enter the number of sheets for any of the following items you are filing with this form.
Do not count copies of the same document

Continuation sheets of this form	
Description	34
Claim(s)	4
Abstract	1
Drawing (s)	3

NIL

10. If you are also filing any of the following, state how many against each items.

Priority documents	
Translations of priority documents	
Statemtent of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature B. A. Yorke & Co. Date **13/05/1996**

B. A. Yorke & Co.

12. Name and daytime telephone number of person to contact in the United Kingdom

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patent Act 1977 stops you from applying for a patent abroad without first getting written permission form the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered "Yes" Patents Form 7/77 will need to be filed.*

e) *Once you have filed in the form you must remember so sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

IMPROVEMENTS IN OR RELATING TO ORGANIC COMPOUNDS

The present invention relates to the production of genetically transformed plants. In particular the invention relates *inter alia* to a process for inducing apomixis, to the apomictic seeds which result from the process, and to the plants and progeny thereof which result from the germination of such seeds.

Apomixis, which is vegetative (non-sexual) reproduction through seeds, is a genetically controlled reproductive mechanism found in some polyploid non-cultivated species. The process is classified as gametophytic or non-gametophytic. In gametophytic apomixis - of which there are two types (apospory and diplospory), multiple embryo sacs which typically lack antipodal nuclei are formed, or else megasporogenesis in the embryo sac takes place. In adventitious embryony (non-gametophytic apomixis), a somatic embryo develops directly from the cells of the embryo sac, ovary wall or integuments. In adventitious embryony, somatic embryos from surrounding cells invade the sexual ovary, one of the somatic embryos out-competes the other somatic embryos and the sexual embryo and utilizes the sexually produced endosperm.

Were apomixis to be a controllable and reproducible phenomenon it would provide many advantages in plant improvement and cultivar development in the case that sexual plants are available as crosses with the apomictic plant.

For example, apomixis would provide for true-breeding, seed propagated hybrids Moreover, apomixis could shorten and simplify the breeding process so that selfing and progeny testing to produce and/or stabilize a desirable gene combination could be eliminated. Apomixis would provide for the use as cultivars of genotypes with unique gene combinations since apomictic genotypes breed true irrespective of heterozygosity. Genes or groups of genes could thus be "pyramided and "fixed" in super genotypes. Every superior apomictic genotype from a sexual-apomictic cross would have the potential to be a cultivar. Apomixis would allow plant breeders to develop cultivars with specific stable traits for such characters as height, seed and forage quality and maturity. Breeders would not be limited in their commercial production of hybrids by (i) a cytoplasmic-nuclear interaction to produce male sterile female parents or (ii) the fertility restoring capacity of a pollinator. Almost all cross-compatible germplasm could be a potential parent to produce apomictic hybrids.

Finally, apomixis would simplify commercial hybrid seed production. In particular, (i) the need for physical isolation of commercial hybrid production fields would be eliminated; (ii) all available land could be used to increase hybrid seed instead of dividing space between pollinators and male sterile lines; and (iii) the need to maintain large quantities of parental line seed stocks would be eliminated.

The potential benefits to accrue from the production of seed *via* apomixis are presently unrealized, to a large extent because of the problem of engineering apomictic capacity into plants of interest. The present invention provides a solution to that problem in that it provides the means for obtaining plants which exhibit the adventitious embryony type of apomixis.

According to the present invention there is provided a method of producing apomictic seeds comprising the steps of:

(i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic,

(ii) regenerating the thus transformed material into plants, or carpel-containing parts thereof, and

(iii) expressing the sequence in the vicinity of the embryo sac.

By "vicinity of the embryo sac" is meant in one or more of the following: carpel, integuments, ovule, ovary wall, chalaza, nucellus, funicle and placenta. The skilled man will recognize that the term "integuments" also includes those tissues, such as endothelium, which are derived therefrom. By "embryogenic" is meant the capability of cells to develop into an embryo under permissive conditions, i.e. in the absence of exogenously added growth regulators such as phytohormones. It will be appreciated that the term "in an active form" includes proteins which are truncated or otherwise mutated with the proviso that they initiate or amplify embryogenesis whether or not in doing this they interact with the signal transduction components that they otherwise would in the tissues in which they are normally present.

The term "plant material" includes protoplasts, isolated plant cells (such as stomatal guard cells) possessing a cell wall, pollen, whole tissues such as emerged radicle, stem, leaf, petal, hypocotyl section, apical meristem, ovaries, zygotic embryo *per se*, roots, vascular bundle, pericycle, anther filament, somatic embryos and the like.

The nucleotide sequence may be introduced into the plant material, *inter alia*, *via* a bacterial or viral vector, by micro-injection, by co-incubation of the plant material and sequence in the presence of a high molecular weight glycol or by coating of the sequence onto the surface of a biologically inert particle which is then introduced into the material.

Expression of the sequence may yield a protein kinase capable of spanning a plant cell membrane. Typically the kinase may be a leucine rich repeat receptor like kinase which has the capacity to auto-phosphorylate. The skilled man will recognize what is meant by the term "leucine rich repeat receptor like kinase". Examples of such proteins include *Arabidopsis* RLK5 (Walker, 1993), *Arabidopsis* RPS2 (Bent *et al.* 1994), Tomato *CF-9* gene product (Jones *et al.* 1994), Tomato N (Whitham *et al.* 1994), *Petunia* PRK1 (Mu *et al.* 1994), the product of the *Drosophila Toll* gene (Hashimoto *et al.* 1988), the protein kinase encoded by the rice *OsPK10* gene (Zhao *et al.* 1994), the translation product of the rice EST clone ric2976 and the product of the *Drosophila Pelle* gene (Shelton and Wasserman, 1993). Still further examples of such proteins include the TMK1, Clavatal, Erecta, and TMKL1 gene products from *Arabidopsis*, the Flightless-1 gene product from *Drosophila*, the TrkC gene product from pig, the rat LhCG receptor and FSH receptor, the dog TSH receptor, and the human Trk receptor kinase. The protein may comprise a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain. In many receptor kinases the extracellular (ligand binding) domain serves as an inhibitor of the kinase domain in the ligand-free state. This arrest is removed after binding of the ligand. Accordingly, in one embodiment of the invention the protein either lacks a ligand binding domain or the domain is functionally inactivated so that the kinase domain can be constitutively active in the absence of an activating signal (ligand). Whether or not the protein possesses a ligand binding domain - functional or otherwise, once expressed and incorporated into the plant cell membrane the protein binding domain is preferably located intra-cellularly.

In a preferred embodiment of the method, the said sequence further encodes a cell membrane targeting sequence. The sequence may be that which is depicted in SEQ ID Nos. 1 or 2 or it may be similar in that it is complementary to a sequence which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity. By "similar" is meant a sequence which is complementary to a test sequence which is capable of hybridizing to the inventive sequence. When the test and inventive sequences are double stranded the nucleic acid constituting the test sequence preferably has a TM within 20°C

of that of the inventive sequence. In the case that the test and inventive sequences are mixed together and denatured simultaneously, the TM values of the sequences are preferably within 10°C of each other. More preferably the hybridization is performed under stringent conditions, with either the test or inventive DNA preferably being supported. Thus either a denatured test or inventive sequence is preferably first bound to a support and hybridization is effected for a specified period of time at a temperature of between 50 and 70°C in double strength citrate buffered saline (SSC) containing 0.1%SDS followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, at a particular temperature, - such as 60°C, for example - such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridization of which is least affected by washing in buffers of reduced concentration. It is most preferred that the test and inventive sequences are so similar that the hybridization between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing 0.1%SDS.

The sequence may be modified in that known mRNA instability motifs or polyadenylation signals may be removed and/or codons which are preferred by the plant into which the sequence is to be inserted may be used so that expression of the thus modified sequence in the said plant may yield substantially similar protein to that obtained by expression of the unmodified sequence in the organism in which the protein is endogenous.

In order to obtain expression of the sequence in the regenerated plant (and in particular the carpel thereof) in a tissue specific manner the sequence is preferably under expression control of an inducible or developmentally regulated promoter, typically one of the following: a promoter which regulates expression of SERK genes *in planta*, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter. The DcEP3-1 gene is expressed transiently during inner integument degradation and later in cells that line the inner part of the developing endosperm. The AtChiIV gene is transiently expressed in the micropylar endosperm up to cellularisation. The LTP-1 promoter is active in the epidermis of the developing nucellus, both integuments, seed coat and early embryo. The bel-1 gene is expressed in the

developing inner integument and the fbb-7 promoter is active during embryo sac development. The *Arabidopsis* ANT gene is expressed during integument development, and the O126 gene from *Phalaenopsis* is expressed in the mature ovule.

It is most preferred that the sequence is expressed in the somatic cells of the embryo sac, ovary wall, nucellus, or integuments.

The endosperm within the apomictic seed results from fusion of polar nuclei within the embryo sac with a pollen-derived male gamete nucleus. It is preferred that the sequence encoding the protein is expressed prior to fusion of the polar nuclei with the male gamete nucleus.

The invention further includes recombinant DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID No. 3, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity. By substantially similar is meant is meant a pure protein having an amino acid sequence which is at least 90% similar to the sequence of the proteins depicted in SEQ ID No 3 below. In the context of the present invention, two amino acid sequences with at least 90% similarity to each other have at least 90% identical or conservatively replaced amino acid residues in a like position when aligned optimally allowing for up to 8 gaps with the *proviso* that in respect of each gap a total not more than 4 amino acid residues is affected. For the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

(i) Serine and Threonine;
(ii) Glutamic acid and Aspartic acid;
(iii) Arginine and Lysine;
(iv) Asparagine and Glutamine;
(v) Isoleucine, Leucine, Valine and Methionine;
(vi) Phenylalanine, Tyrosine and Tryptophan
(vii) Alanine and Glycine

It is preferred that the DNA further encodes a cell membrane targeting sequence, and that the protein encoding region is under expression control of a developmentally regulated or inducible promoter, such as, for example, a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene

promoter, the *Arabidopsis* ANT gene promoter, or the promoter of the O126 gene from *Phalaenopsis*.

Particularly preferred embodiments of the said DNA include those depicted in SEQ ID Nos. 1 or 2 or those which are complementary to one which hybridizes under stringent conditions with the said sequences and which encode a membrane bound protein having kinase activity. As indicated above, the DNA may be modified in that known mRNA instability motifs or polyadenylation signals may be removed and/or codons which are preferred by the plant into which the DNA is to be inserted may be used so that expression of the thus modified DNA in the said plant may yield substantially similar protein to that obtained by expression of the unmodified DNA in the organism in which the protein is endogenous.

The invention still further includes a vector which contains DNA as indicated in the three immediately preceding paragraphs, plants transformed with the recombinant DNA or vector, and the progeny of such plants which contain the DNA stably incorporated, and/or the apomictic seeds of such plants or such progeny. Such plants include field crops, vegetables and fruits including tomato, pepper, melon, lettuce, cauliflower, broccoli, cabbage, brussels sprout, sugar beet, corn, sweetcorn, onion, carrot, leek, cucumber, tobacco, alfalfa, aubergine, beet, broad bean, celery, chicory, cow pea, endive, gourd, groundnut, papaya, pea, peanut, pineapple, potato, safflower, snap bean, soybean, spinach, squashes, sunflower, sorghum, water-melon and the like; and ornamental crops including Impatiens, Begonia, Petunia, Pelargonium, Viola, Cyclamen, Verbena, Vinca, Tagetes, Primula, Saint Paulia, Ageratum, Amaranthus, Anthirrhinum, Aquilegia, Chrysanthemum, Cineraria, Clover, Cosmo, Cowpea, Dahlia, Datura, Delphinium, Gerbera, Gladiolus, Gloxinia, Hippeastrum, Mesembryanthemum, Salpiglossis, Zinnia, and the like. In a preferred embodiment, the DNA is expressed in "seed crops" such as corn, sweet corn and peas etc. in such a way that the apomictic seed which results from such expression is not physically mutated or otherwise damaged in comparison with seed from untransformed like crops.

The invention still further includes the use of the said DNA in the manufacture of apomictic seeds, in which use the sequence is expressed in the vicinity of the embryo sac.

The invention will be further apparent from the following description and the associated drawings and sequence listings.

SEQ ID No. 1 depicts the *Daucus carota* genomic clone of the putative receptor kinase (SERK) associated with the transition of competent to embryogenic cells; SEQ ID No. 2 depicts the cDNA of the said putative kinase; SEQ ID Nos. 3-16 depict the sequences of various PCR primers; and SEQ ID Nos. 17-19 depict specific peptides contained within the gene product of SEQ ID No. 2.

Figure 1 shows the results of an RT-PCR experiment performed on RNA extracted from the indicated tissues. 40 cycles followed by Southern blotting of the resulting bands is necessary to visualize SERK expression. Lanes include explants at day 7, treated for less (lane 1) or more (lane 2) then 3 days with 2,4-D. In the original a very faint signal is visible in lane 2, but not in lane 1. Established embryogenic cultures (lanes 4-6) but not a non-embryogenic control (lane 3) express the SERK gene. In carrot plants, no expression is detectable except for developing seeds after pollination (lane 7). Up to day 7 after pollination, the carrot zygote remains undivided, suggesting that the observed signal is coming only from the zygote. At day 10, the early globular and at day 20 the heart stage is reached in carrot zygotic embryogenesis. No signals are seen on Northern blots.

Figure 2A shows the results of a whole-mount *in situ* hybridization with the SERK cDNA on 7 day explants treated for 3 days with 2,4 D. Few cells on the surface of the explant express the SERK gene, and those cells that do are the ones that become embryogenic. Figure 2B shows a whole mount *in situ* hybridization on a partially dissected seed containing a globular zygotic embryo. Hybridization is visualized by DIG staining.

Figure 3 shows SERK expression in embryogenic hypocotyl cells during hormone-induced activation, determined by whole mount *in situ* hybridization . Bar: 50 µm
(A-E) Cell population generated by mechanical fragmentation of the activated hypocotyls. Only few of a certain type of cell, defined enlarged cell show SERK expression (asterisks). Small cytoplasmic cells (c), enlarging cells (eg) and large cells (l) never show SERK expression.
(F) Hypocotyl longitudinal section before hormone-induced activation. It is not possible to detect any SERK expression in any type of cell.
(G-I) Proliferating mass coming from the inner hypocotyl tissues 10 days after the beginning of the hormonal treatment (longitudinal section). In G a single enlarged cells showing SERK expression is detectable within a row of negative cells showing the same morphology. In H a single enlarged cell showing serk expression is detaching from the surface of the proliferating

mass. In I a cluster of enlarged cells showing SERK expression is detectable at the surface of proliferating tissue.

(J) Proliferating mass coming from the inner tissues of the hypocotyl 10 days after the beginning of the rooting treatment (24 hours with 2,4-D followed by hormone removal). Both the root primordia and the enlarged cells detaching from the surface do not show any SERK expression.

The following description illustrates the isolation and cloning of the SERK gene and the production of apomictic seed by heterologous expression of the said gene in the nucellar region of the carpel so that somatic embryos form which penetrate the embryo sac and are encapsulated by the seed as it develops.

ISOLATION AND CLONING OF THE SERK GENE FROM *DAUCUS CARROTA*

Isolation of cDNA clones that are preferentially expressed in embryogenic cell cultures of carrot

In order to increase the chance of success for obtaining genes expressed in carrot suspension cells competent to form embryos, the number of embryo-forming cells as present in a series of established cell cultures was determined. A sub-population of cells that passed through a 30 µm nylon sieve was isolated from eight different cultures that ranged in age between 2 months and 4 years. In these sub 30 µm populations, the number of embryos formed from the single cells and small cell clusters was determined and expressed as a percentage of the total number of cells present at the start of embryogenesis. Sieved <30 µm cultures able to form somatic embryos with a frequency of more than 1% were then used as a source for competent cells, and cultures that produced less than 0.01% embryos were used as non-embryogenic controls. As main cloning strategies, cold plaque screening (Hodge *et al.* 1992) and differential display (dd) RT-PCR (Liang and Pardee, 1992) were used besides conventional differential screening of cDNA libraries.

Labeled probes for differential screening were obtained from RNA out of a <30 µm sieved sub-population of cells from either embryogenic or non-embryogenic cell cultures. Employing these probes in a library screen of approximately 2000 plaques yielded 26 plaques that failed to show any hybridization to either probe. These so-called cold plaques were purified and used for further analysis. From the total number of plaques that did hybridize, about 30 did so only with

the probe from embryogenic cells. ddRT-PCR reactions using a combination of one anchor primer and one decamer primer were performed on mRNA isolated from three embryogenic, and three non-embryogenic suspension cultures. About 50 different ddRT-PCR fragments were obtained from each reaction. Using combinations of three different anchor and six different decamer primers, a total of approximately 1000 different cDNA fragments was visualized. Six of these PCR fragments were only found in lanes made with mRNA from <30 µm populations of cells from embryogenic cultures (Table 1) and with oligo combinations of the anchor primer (5'-TTTTTTTTTTTGC-3') and the decamer primers (5'-GGGATCTAAG-3'), (5'-ACACGTGGTC-3'), (5'-TCAGCACAGG-3'). Because differential PCR fragments often consist of several unresolved cDNA fragments (Li *et al.* 1994), cloning proved to be essential prior to undertaking further characterization of the PCR fragments obtained.

All clones obtained were subjected to a second screen, that consisted of spot-dot Northern hybridization performed under conditions of high stringency. This method, that used RNA from entire unsieved embryogenic and non-embryogenic suspension cultures, proved to be a fast and reliable additional selection method. Only one clone (22-28) of the 30 clones obtained after differential screening, proved to be restricted to embryogenic cell cultures while the majority was constitutively expressed. The 26 clones obtained from the cold plaque screening required long exposure times in the spot-dot Northern analysis. Six of these clones failed to show any hybridization signal and 19 proved to be expressed in both embryogenic and non-embryogenic cell cultures. One clone (31-50) showed low expression in all embryogenic cultures, and in one non-embryogenic culture, but not in the others. Of the six cloned fragments obtained by ddRT-PCR display, four showed hybridization more or less restricted to transcripts present in embryogenic cultures. All clones that passed through the second screening were sequenced. Two of the ddRT-PCR clones (6-8 and 7-13) were identical to the carrot Lipid Transfer Protein (LTP) gene, previously identified as a marker for embryogenic carrot cell cultures. LTP expression is restricted to embryogenic cell clusters and the protoderm of somatic and zygotic embryos from the early globular stage onwards (Sterk *et al.* 1991). Therefore, while the LTP gene is not a marker for competent cells, its appearance in the screening confirms the validity of our methods with respect to the cloning of genes expressed early during somatic embryogenesis.

cDNA clone 31-50 encodes a leucine-rich repeat containing receptor-like kinase

The mRNA corresponding to the isolated clone 31-50 had an open reading frame of 1659 nucleotides encoding a protein with a calculated Mw of 55 kDa. Because clone 31-50 is mainly

expressed in embryogenic cell cultures it was renamed Somatic Embryogenesis Receptor Kinase (SERK). The SERK protein contains a N-terminal domain with a five-times repeated leucine-rich motif that is proposed to act as a protein-binding region in LRR receptor kinases (Kobe and Deisenhofer, 1994). Between the extracellular LRR domain of SERK and the membrane-spanning region is a 33 amino acid region rich in prolines (13), that is unique for the SERK protein. Of particular interest is the sequence SPPPP, that is conserved in extensins, a class of universal plant cell wall proteins (Varner and Lin, 1989). The proposed intracellular domain of the protein contains the 11 subdomains characteristic of the catalytic core of protein kinases. The core sequences HRDVKAAN and GTLGYIAPE in respectively the kinase subdomains VB and VIII suggest a function as a serine / threonine kinase (Hanks *et al.* 1988). Another interesting feature of the intracellular part of the SERK protein is that the C-terminal 24 amino-acids resembles a single LRR. The serine and threonine residues present within the intracellular LRR sequence are surrounded by acidic residues and might be targets for the autophosphorylation of SERK, thereby regulating the ability of other proteins to interact with this receptor-kinase in a similar fashion as described for the SH2 domain of the EGF family of tyrosine receptor kinases.

Hybridization of the SERK cDNA clone to the carrot genome revealed the presence of only a single main hybridizing band after digestion with EcoR1, probably reflecting a single SERK gene in the carrot genome. This was confirmed after digestion with DdeI, an enzyme that cuts three times within the SERK gene. No signal was observed after Northern blotting of mRNA from embryogenic cell cultures and hybridization with labeled SERK probes, reflecting the low levels of transcript present in these cultures. Detection of the SERK transcript on the original spot-dot Northerns was only possible after long exposure times compared with other probes.

The ability of the SERK protein to autophosphorylate was investigated *in vitro,* using a previously described autophosphorylation assay (Mu et al. 1994), with a bacterial fusion protein that contained the complete intracellular region of the SERK protein. The bacterially expressed SERK fusion protein was able to autophosphorylate, indicating that the SERK protein is able to fulfill a role as a protein kinase *in vivo* (Heldin, 1995).

Expression of the SERK gene corresponds with the first appearance of competent cells during hypocotyl activation

When carrot hypocotyls are induced with 2,4-D, only the cells of the provascular tissue

proliferate. Cells of epidermal and cortical origin merely expand, suggesting that the provascular tissue derived cells form the newly initiated suspension culture. After removal of 2,4-D, the formation of somatic embryos occurs after 2-3 weeks. Somatic embryos are preceded by embryogenic cells, that are developed in turn from competent cells. While competent and embryogenic cell formation take place in the presence of 2,4-D, it was not clear when this occurred, and which cells acquired competence. Since previous experiments (Toonen et al. 1994) revealed that cell morphology is not a good criterion, the first appearance of single competent cells was determined experimentally by semi-automatic cell tracking performed on large populations of immobilized cells. Hypocotyl explants activated with 2,4-D for seven days were mechanically fragmented and samples of the resulting population of mainly single suspension cells were immobilized to allow recording of their development by cell tracking. In the immobilized cell populations obtained in this way all the morphologically discernible cell types were present that were also seen in the un-fragmented activated hypocotyls. Because the different cell morphologies observed during hypocotyl activation were known (Guzzo *et al.* 1995), it was possible to trace back the original position of each type of cell in the activated explant. Small cytoplasm-rich cells (16x16 μm) are the proliferating cells that surround the vascular elements. Enlarging vacuolated cells (16x40 μm) are encountered on the surface of the mass of proliferating cells and these can detach from the surface when fully enlarged (35x90 μm). Large vacuolated cells (more than 60x140 μm) are the non-proliferating remnants of the hypocotyl epidermis and cortical parenchyma. The shape of the enlarging and fully enlarged cells could change from oval to elongate or triangular. Cell tracking on a total of 24,722 cells released from seven days activated hypocotyls showed that only 20 single cells formed a somatic embryo. Because of their dependance on continued 2,4-D treatment, the embryo-forming single cells are still in the competent cell stage. All of the embryo-forming single cells belonged to the category of 3,511 enlarged cells, that contained therefore competent cells in a frequency of 0.56%. The single cell tracking experiments clearly reveal that the ability of explant cells to reinitiate cell division under the influence of 2,4-D, resulting in a population of highly cytoplasmic and rapidly proliferating cells, does have a causal relation with the ability to become embryogenic. It is also clear that only a very limited number of the cells that make up the newly initiated embryogenic suspension culture are actually competent to form embryogenic cells.

Expression of the SERK gene, determined by whole mount *in situ* hybridization on a similar population of cells as used for the cell tracking experiments, was found to be restricted to only 0.44% of the enlarged cells. Therefore, the expression of the SERK gene appears closely

correlated both qualitatively and quantitatively with the presence of competent single cells.

To obtain insight into the temporal regulation of SERK expression in the course of explant activation, whole mount *in situ* hybridization was performed on entire intact or hand-sectioned explants treated for different periods with 2,4-D. Representative samples were collected at daily intervals from explants untreated and treated for three days, six days, seven days or ten days with 2,4-D before returning to B5-0. No SERK-expressing cells were ever found in explants treated for less then three days with 2,4-D. While enlarged cells became present after the first five days of culture, the first few SERK-expressing enlarged cells were found after six-seven days of culture in the presence of 2,4-D treatment. These few cells were present at the surface of the mass of proliferating cells originating from the provascular tissue. In the hypocotyls treated for ten days with 2,4-D, the number of SERK-positive cells had increased to 3.04% and included at this stage also cells present in small clusters. No SERK transcript was ever detected in small cytoplasm-rich cells or large vacuolated cells. Hypocotyls were also treated for only one day with 2,4-D and subsequently cultured in hormone-free medium for a total of seven or ten days. Under these conditions explant cells proliferated and gave rise to roots and non-embryogenic cell cultures, while SERK expression could never be detected. The in situ hybridization results described above were obtained from a relatively small number of explants and a few hundred cells, so RT-PCR followed by Southern hybridization was performed to obtain more quantitative results. These are shown in Figure 7 and confirm the close temporal correlation between the first appearance of competent cells in explants treated for three days with 2,4-D and the expression of the SERK gene. Northern hybridization never gave any signal after hybridization with SERK cDNA probes, not even after prolonged exposure in a PhosphorImager, in line with the extremely restricted expression pattern of the SERK gene.

Expression of the SERK gene corresponds with the occurrence of competent cells in established embryogenic cell cultures

While the results described so far indicate that competent and embryogenic cell formation is restricted to a particular class of enlarged cells during explant activation, the situation in an established embryogenic cell culture is more complex. Competent single cells in such cultures do not appear to belong to one cell type in particular, but have been shown to originate from all morphologically different cell types. In cell tracking experiments, embryogenic cells, that do not require exogenous auxin treatment, were never observed to be single but consisted of clusters of at least 3-4 cells (Toonen et al. 1994). SERK expression was found in all morphologically

discernible single cell types that were present in an embryogenic cell culture at a frequency between 0.1 and 0.5% depending on the cell_type. In non-embryogenic cultures, SERK expressing cells were never encountered. As was observed in the activated explants, SERK expression was not restricted to single cells, but also occurred in small clusters of 2 to 16 cells. Since clusters of this size are known to consist of embryogenic cells, these data show that SERK expression is not restricted to competent single cells, but may persist in small clusters of embryogenic cells. No SERK expression was encountered during the late globular, heart and torpedo-stages of somatic embryogenesis.

The SERK gene is transiently expressed in zygotic embryogenesis

The expression of the SERK gene in carrot plants was determined by RT-PCR. The results indicate that no SERK mRNA accumulates in any of the adult plant organs nor in flowers prior to pollination. The first occasion when SERK expression can be detected is in flowers at three days after pollination (DAP), at which stage fertilization has taken place and endosperm development has commenced. SERK mRNA remains present in flowers up to twenty DAP, corresponding with the early globular stage of the zygotic embryo (Yeung et al. 1996). Whole mount *in situ* hybridization on partially dissected carrot seeds confirmed that the SERK gene was only expressed in early embryos up to the globular stage. Expression was observed in the entire embryo including the suspensor. No expression was seen in seedlings, roots, stems, leaves, developing and mature flower organs, pollen grains and stigma's before and after fertilization. Tissues in the developing seed such as seed coat, integuments, all embryo sac constituents before fertilization as well as the endosperm at all stages of development investigated did not show any SERK expression. Later stages of carrot zygotic embryos were also completely devoid of SERK mRNA. Given this pattern of expression, that is restricted to the zygotic embryo, the signal as detected by RT-PCR in flowers at 3 and 7 DAP must come from SERK mRNA as present in zygotes, because in carrot the zygote remains undivided up to one week after pollination (Yeung et al. 1996). Although SERK expression persists to slightly later stages in zygotic globular embryos when compared to the somatic ones, these results confirm the transient pattern of expression as observed for the SERK gene during somatic embryogenesis and also imply that there is a correspondence between the formation of competent cells in vitro and the formation of the zygote in vivo.

METHODS

Cell culture, hypocotyl explant induction and cell tracking

Cell cultures were derived from *Daucus carota*_cv. Flakkese and maintained as previously described (De Vries *et al.* 1988a). Cell suspension cultures were maintained at high cell density in B5 medium (Gamborg *et al.* 1968) supplemented with 2 µM 2,4-D (B5-2 medium). Embryo cultures with globular, heart and torpedo-stage somatic embryos were derived from <30 µm sieved cell cultures cultured at low cell density (100 000 cells / ml) in B5 medium without 2,4-D (B5-0). For hypocotyl explant induction experiments, plantlets were obtained from seed of *Daucus carota* cv. S Valery as described previously (Guzzo *et al.*, 1994). The hypocotyls of one week old plantlets were divided in segments of 3-5 mm, incubated for various periods of time in B5-2 medium and returned to B5-0 medium. Seven days after explantation and exposure to 2,4-D the hypocotyl segments were fragmented on a 170 µm sieve and the resulting cells collected to form a fine cell suspension. Immobilization of these cells in B5-0.2 medium was performed in a thin layer of phytagel (Toonen *et al.* 1994). After one week of further culture 2,4-D was removed by washing the plates with B5-0 medium. This allowed embryos to develop beyond the globular stage. Recording the development of the immobilized cells was performed with a procedure modified from the previously described by Toonen et al. (1994). The main change involved a new MicroScan program for automatic 3-axis movement to scan all cells in the phytagel (Toonen et al. 1996).

Nucleic acid isolation and analysis

RNA was isolated from cultured cells and plant tissues as described by De Vries *et al.* (1988b). Poly(A)$^+$-RNA was obtained by purification by oligo (dT) cellulose (Biolabs). For RNA gel blot analysis samples of 10 µg total RNA were electrophoresed on formamide gel, and transferred to nytran-plus membranes. For RNA spot-blot analysis 5 µg of total RNA was denatured and spotted onto nytran-plus filters using a hybridot manifold (BRL).

Genomic DNA was isolated according to Sterk et al. (1991). Samples of 10 µg genomic DNA were digested with different restriction enzymes and separated on agarose gel, and transferred to nytran-plus membrane (Schleicher & Schuell). Hybridization of RNA blots took place at 42°C in hybridization buffer containing 50% formamide, 6xSSC, 5xDenhardt, 0.5% SDS and 0.1 mg/ml salm sperm DNA. Hybridization of DNA blots was performed as previously described (Sterk et al. 1991). Following hybridization, filters were washed under stringent conditions (3x20 min in 0.1% SSC, 1% SDS, at 65°C). Filters were exposed to Kodak X-Omat AR film. The integrity and the amount of RNA on the blots was confirmed by hybridization with an 18S

ribosomal RNA probe. Nucleotide sequence analysis was performed on an ABI 373A automated DNA sequencer (Applied Biosystem).

Screening procedures

Two independent cDNA libraries were constructed with equal amounts of poly(A)$^+$-RNA from total established cell cultures grown for six days in B5-2 medium, sieved <125 µm cell cultures grown for six days in B5-0 medium and sieved <30 µm cell cultures grown for six days in B5-0 medium. cDNA synthesis and cloning into the Uni-ZAP™ XR vector was performed according to the manufacturers protocol (Stratagene).

Differential screening of the cDNA libraries was performed essentially as described by Scott *et al.* (1991). RNA was isolated from either three embryogenic or three non-embryogenic cell cultures, that were grown for seven days in B5-2 after sieving through 30 µm mesh. First strand cDNA synthesis was performed on 4 µg total RNA using AMV reverse transcriptase (Gibco BRL). [^{32}P]dATP labeled probes were prepared using random prime labeling on first strand cDNA. Pooled probes from embryogenic and non-embryogenic cell populations were hybridized to two pairs of nitrocellulose filters, each containing 1000 plaques from one cDNA library. After washing for 3x20 min in 0.1% SSC, 1% SDS at 65°C, hybridization was visualized by autoradiography for two days on Kodak X-omatic film. Plaques that only showed signal with the embryogenic transcript probe were purified by two further rounds of screening.

In order to identify cDNA clones which are expressed at low levels in the <30 µm sieved cell population, cold plaque screening was performed as described by Hodge *et al.* (1992). Plaques from the differential screening that did not show any signal after seven days of autoradiography were purified by two further rounds of screening. The resulting clones were used as probes for characterization of the expression pattern of the corresponding genes.

Differential Display RT-PCR

Differential display of mRNA was performed essentially as described by Liang and Pardee (1992). cDNA synthesis took place by annealing 1 µg of total RNA in 10 µl buffer containing 200 mM KCl, 10 mM Tris-HCl (pH 8.3), and 1 mM EDTA with 100 ng of one of the following anchor primers: (5'-TTTTTTTTTTTGC-3'), (5'-TTTTTTTTTTTCTG-3'), (5'-TTTTTTTTTTTCA-3'). Annealing took place by heating the mix for 3 min. at 83°C followed by incubation for 30 min at 42°C. Annealing was followed by the addition of 15 µl pre-warmed

cDNA buffer containing 16 mM $MgCl_2$, 24 mM Tris-HCl (pH 8.3), 8 mM DTT, 400 µM dNTP, and 4 Units AMV reverse transcriptase (Gibco BRL). cDNA synthesis took place at 42°C for 90 min. First strand cDNA was phenol/chlorophorm extracted and precipitated with ethanol using glycogen as a carrier. The PCR reaction was performed in a reaction volume of 20 µl containing 10% of the synthesized cDNA, 100 ng of anchor primer, 20 ng of one of the following 10-mer primers: (5'-GGGATCTAAG-3') , (5'-TCAGCACAGG-3'), (5'-GACATCGTCC-3'), (5'-CCCTACTGGT-3'), (5'-ACACGTGGTC-3'), (5'-GGTGACTGTC-3'), 2 µM dNTP, 0.5 Unit*Taq* enzyme in PCR buffer (10 mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin and 0.1% Triton X100) and 6 nM [a-^{32}P] dATP (Amersham). PCR parameters were 94°C for 30 sec, 40°C for 1 min, and 72°C for 30 sec for 40 cycles using a Cetus 9600 (Perkin-Elmer). Amplified and labeled cDNAs were separated on a 6% denaturing DNA sequencing gel. Gels were dried without fixation and bands were visualized by 16 hours of autoradiography using Kodak X-omatic film. Bands containing differentially expressed cDNA fragments of 150-450 nucleotides were cut out of the gel and DNA was extracted from the gel slices by electroelution onto DE-81 paper (Whatmann). After washing of the paper in low salt buffer (100 mM $LiCl_2$ in 10 mM TE buffer), and elution of the cDNA in high salt buffer (1 M $LiCl_2$ in 10 mM TE buffer with 20% ethanol) the cDNA was concentrated by precipitation in ethanol using glycogen as carrier. Reamplification of the cDNA fragments using the same PCR cycling parameters as described above but PCR buffer containing 2.5 µM of both the 10-mer and the anchor oligo and 100 µM dNTP. DE-81 paper allowed an efficient recovery of the DNA fragments and reamplification generated an average of 500 ng DNA after 40 cycles. Amplified PCR products were blunt-ended using the Klenow fragment of *E.coli* DNA Polymerase I (Pharmacia), purified on Sephacryl-S200 columns (Pharmacia), ligated into a SmaI linearized pBluescript vector II SK$^-$ (Stratagene) and transformed into *E.coli* using electroporation.

RT-PCR

Adult plant tissues from *Daucus carota* were obtained from S&G Seeds (Enkhuizen). Controlled pollination was performed by hand. Flower tissue RNA was obtained from three compete umbels for each time-point and contained all flower organs including pollen grains. 2 µg of total RNA from adult plant tissue or cell cultures was annealed at 42°C with 50 ng oligo (5'-TCTTGGACCAGATAATTC-3') in 10 µl annealing buffer (250 mM KCl, 10 mM Tris-HCl pH 8.3, 1 mM EDTA). After 30 min. annealing, 1 unit AMV-reverse transcriptase was added in a volume of 15 µl cDNA buffer (24 mM Tris-HCl pH 8.3, 16 mM $MgCl_2$, 8 mM DTT, 0.4 mM dNTP). The reverse transcription reaction took place for 90 min. at 42°C. PCR amplification of

SERK-cDNA was carried out with two specific oligos for the SERK kinase domain, (5'-CTCTGATGACTTTCCAGTC-3') and (5'-AATGGCATTTGCATGG-3'). Amplification was carried out with 30 cycles of 30 sec. at 94°C, annealing at 54°C for 30 sec. and extension at 72°C for 1 min., followed by a final extension for 10 min.at 72°C.

Whole mount *in situ* hybridization

Whole mount *in situ* hybridizations were performed essentially as previously described (Engler et al. 1994). Cell cultures and somatic embryos were immobilized on poly-L-lysine coated glasses during fixation to improve handling. Whole mount *in situ* hybridization on explants took place by embedding hypocotyls from seven-days old plantlets in 3% Seaplaque agarose (Duchefa) and processing them in Eppendorf tubes. Transverse as well as longitudinal sections were made with a vibrotome (Biorad Microcut). Sections of 50-170 µm thick were incubated in B5-2 medium for a minimum of three days to induce formation of embryo-forming cells. Optimal induction was achieved with longitudinal hypocotyl sections with a thickness of at least 90 µm. To obtain proliferating, non-embryogenic cell cultures, hypocotyl sections were exposed to 2,4-D for only 1 day, and subsequently transferred to B5-0 medium (Guzzo *et al.* 1994). Whole mount in situ hybridization on developing seeds was performed by removing the chalazal end of the seeds to allow easier probe penetration. After hybridization, the enveloping layers of integuments and endosperm were carefully removed to expose the developing embryos. In situ hybridization on sections was performed as described previously (Sterk et al. 1991) except for the use of non-radioactive probes.

All samples were fixed for 60 min. in PBS containing 70 mM EGTA, 4% paraformaldehyde, 0.25% glutaraldehyde, 0.1% Tween 20, and 10% DMSO. Samples were then washed, treated with proteinase K for 10 min, again washed and fixed a second time. Hybridization solution consisted of PBS containing 0.1% Tween 20, 330 mM NaCl, 50 µg/ml heparin, and 50% deionized formamide. Hybridization took place for 16 hours at 42°C using digoxigenin-labeled sense or antisense riboprobes (Boehringer Mannheim). After washing the cells were treated with RNaseA, and incubated with anti-digoxigenin-alkaline phosphatase conjugate (Boehringer Mannheim) which had been preabsorbed with a plant protein extract. Excess antibody was removed by washing followed by rinsing in staining buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM levamisole) and the staining reaction was performed for 16 hours in a buffer containing NBT and BCIP. Observations were performed using a Nikon Optiphot microscope equipped with Nomarski optics.

Autophosphorylation assay

A 1.4 kB SspI cDNA fragment of the SERK cDNA encoding most of the open reading frame apart from the N-terminal three LRRs was cloned into the pGEX expression vector (Pharmacia). A fusion protein consisting of SERK and the glutathione S-transferase gene product was synthesized by a three hours induction of transformed *E.coli* with 2 mM IPTG. Fusion protein was isolated and purified as described previously (Horn and Walker, 1994). Purified fusion protein was coupled to glutathione agarose beads (Sigma) and incubated for 20 min. at 20°C in a volume of 10 µl buffer: 50 mM Hepes (pH 7.6), 10 mM $MgCl_2$, 10 mM $MnCl_2$, 1 mM DTT, 1 µCi [γ -^{32}P] (3 000 Ci/mmol) . Excess label was removed by washing the fusion protein/glutatione agarose beads three times for 5 min. in 50 mM Tris-HCl (pH 7.3), 10 mM $MgCl_2$ at 4°C. Protein was removed from the beads by cooking in SDS-PAGE loading buffer. Equal amounts of protein were separated by SDS-PAGE and protein autophosphorylation was visualized by autoradiography.

INTRODUCTION OF THE SERK GENE INTO *PLANTA*
AND THE PRODUCTION OF APOMICTIC SEED

The promoters of the DcEP3-1 and the AtChit IV genes are cloned and characterized by standard procedures. The DcSERK coding sequence (SEQ ID No. 2) is cloned behind the DcEP3-1, the AtChit IV or the AtLTP-1 promoters and transformed into *Arabidopsis*. The ligation is performed in such a way that the promoter is operably linked to the sequence to be transcribed. This construct, which also contains known marker genes providing for selection of transformed material, is inserted into the T-DNA region of a binary vector such as pBIN19 and transformed into *Arabidopsis*. *Agrobacterium*-mediated transformation into *Arabidopsis* is performed by the vacuum infiltration or root transformation procedures known to the skilled man. Transformed seeds are selected and harvested and (where possible) transformed lines are established by normal selfing. Parallel transformations with 35S promoter-SERK constructs and the entire SERK gene itself are used as controls to evaluate over-expression in many cells or only in the few cells that naturally express the SERK gene. The 35S promoter-SERK construct may give embryo formation wherever the signal that activates the SERK-mediated transduction chain is present in the plant. A testing system based on emasculation and the generation of donor plant lines for pollen carrying LTP1 promoter-GUS and SERK promoter-barnase is established.

The same constructs (35S, EP3-1, AtChitIV, AtLTP-1 and SERK promoters fused to the SERK coding sequence) are employed for transformation into several *Arabidopsis* backgrounds. These

backgrounds are wild type, male sterile, fis (allelic to emb 173) and primordia timing (pt)-1 lines, or a combination of two or several of these backgrounds. The wt lines are used as a control to evaluate possible effects on normal zygotic embryogenesis, and to score for seed set without fertilization after emasculation. The ms lines are used to score directly for seed set without fertilization. The fis lines exhibit a certain degree of seed and embryo development without fertilization, so may be expected to have a natural tendency for apomictic embryogenesis, which may be enhanced by the presence of the SERK constructs. The pt-1 line has superior regenerative capabilities and has been used to initiate the first stably embryogenic *Arabidopsis* cell suspension cultures. Combinations of several of the above backgrounds are obtained by crossing with each other and with lines containing ectopic SERK expressing constructs. Except for the ms lines, propagation can proceed by normal selfing, and analysis of apomictic traits following emasculation. A similar strategy is followed in which the ATChiIV, AtLTP-1 and SERK promoters are replaced by the bel-1 and fbp-7 promoters as well by other promoters specific for components of the female gametophyte.

Additional constructs are generated that have constitutive receptor kinase activity. Most of the receptor kinases of the SERK type act as homodimeric receptors, requiring autophosphorylation before being able to activate downstream signal transduction cascades. In many receptor kinases the extracellular domain serves as an inhibitor of the kinase domain in the ligand-free stage. This arrest is removed after binding of the ligand (Cadena and Gill, 1992). By introduction of a SERK construct, from which the extracellular ligand-binding domain has been removed, mutant homodimeric (in cells that do not have a natural population of SERK proteins) or heterodimeric (in cells that also express the unmodified forms) proteins can be generated with a constitutively activated kinase domain. This approach, when coupled to one of the promoters active in the nucellar region, results in activation of the embryogenic pathway in the absence of the activating signal. This may be an important alternative in cases where it is necessary or desirable to have activation of the SERK pathway only dependant on specific promoter activity and independent of temporal regulation of an activating signal. Introduction of SERK constructs that result in fertilization-independent-embryogenesis (fie) are tested in other species for their effect. In order to recognize the fie phenotype, the skilled man will use appropriate male sterile backgrounds. However, pollination is often necessary for apomixis of the adventitious embryony type, in order to ensure the production of endosperm.

Whilst the present invention has been particularly described by way of the production of

apomictic seed by heterologous expression of the SERK gene in the nucellar region of the carpel, the skilled man will recognize that other genes, the products of which have a similar structure/function to the SERK gene product, may likewise be expressed with similar results. Moreover, although the example illustrates apomictic seed production in *Arabidopsis*, the invention is, of course, not limited to the expression of apomictic seed-inducing genes solely in this plant. Moreover, the present disclosure also includes the possibility of expressing the SERK (or related) gene sequences in the transformed plant material in a constitutive - tissue non-specific manner (for example under transcriptional control of a CaMV35S or NOS promoter). In this case, tissue specificity is assured by the localized presence within the vicinity of the embryo sac of the ligand of the product of the said gene. Furthermore, the SERK (or related) gene products may interact with proteins such as transcription factors which are involved in regulating embryogenesis. This interaction within tissue which has been transformed according to the present disclosure is also part of the present invention.

The skilled man who has the benefit of the present disclosure will also recognize that the SERK gene (and others as indicated in the preceding paragraph) may be transformed into plant material which may be propagated and/or differentiated and used as an explant from which somatic embryos can be obtained. Expression of such sequences in the transformed tissue (which is subjected to a ligand of the kinase gene products) substantially increases the percentage of the cells in the tissue which are competent to form somatic embryos, in comparison with the number present in non-transformed like tissue.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: SANDOZ LTD
(B) STREET: Lichtstrasse 85
(C) CITY: Basel
(D) STATE: BS
(E) COUNTRY: Switzerland
(F) POSTAL CODE (ZIP): CH-4002
(G) TELEPHONE: 061-324-4203
(H) TELEFAX: 061-322-7532

(ii) TITLE OF INVENTION: Improvements in or relating to organic compounds

(iii) NUMBER OF SEQUENCES: 19

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6695 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3696..6617

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 3731..3802

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 3851..3979

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4124..4211

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4284..4357

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4430..4528

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4642..4757

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4890..4967

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 5295..5803

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 6197..6339

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCTAGATGAC GAAATCGCGC TACCTTTGAT TTNGAAATAC TAGGTTGTAG TATCTTGATT 60

AGTTTTTTGG ATATCTTGCT GTAATTTCTT TAGGAGATGC AAACGGTCTT CATTTAATAT 120

GAGCCCTTGT GACTTGACAA AAGTATCTAG CATGTTTGAT CACGAGGTAG CTAAAAAGTA 180

GCGTGTTTGA TTAAGCACAT AATATTGTAT TGGGCCTATT GGCTATCAAT GAAGTTTGAT 240

GCAAGTATAT AGCTTGTATT ATGCATGTGA TGAGGGTATA TAAAAGAAGT AAAGAACATT 300

CTCTCGTAGC ATTCATTTTT CTCTTGCCTA TAGTTAACGA GTTTTGTCAC ACATGACGTT 360

GAAACTGGAT GTGTCTGTTC TTCCATCTAA GTTTGGATTA CCTGATAGAT GCTCAACTTC 420

TTCGTCAGCC TTTTCTTTCC GATTTTTCCC AAGACAAGAT TCTTTAGTTA ATAGTTATTG 480

CTCTGGTGGC TTGTGTGCAT TTTAGGAATC TTACTCTGTT TTTTAATGGA GAAACGAAAC 540

CTACCTTTTT TTCTGTGTTC CCTTTTATGA TATCACCTGC TTGGAGGCGT TTAGACTTTA 600

TCCACCTAAA CTATTCATGT TTACCAGACA AGCTATACGT TTTATCCCCC CCCCCCGCGG 660

ACCTGNGGAC AAAAGAAGCG CTGATGAACT GATTTAATCC GTGTTTTATT ATATTACACA 720

TTGATGCTTC ATGGAGCTAA TATCTTTGGT TAAATTTCAT GTATATATAT ACCCTTCCCT 780

CTTGTGATGG CAGTGGCCCC TCGTTTAATT AGCGTACTTA ATTATCTGAT GGATACTGTA 840

TGCTTGGCAG ATGATGTCAT CAGATTATAC CATTTGTTGT GCTCTACAAA ATAAAAAACC 900

TCTATTTATG TTCATCTTTT TGGTAACAAG TAACTAATTG ATGCGCTATG TTGACAGGCG 960

ATGCATTACA CAACTTACGA ACTAGCTTGC AAGATCCCAA CAATGTCCTG CAGAGCTGGG 1020

ATCCAACCCT TGTGAACCCT TGCACATGGT TTCATGTGAC ATGTAACAAT GAAAACAGTG 1080

TTATAAGAGT GTAGGTCACT TCCCTTATTA ATTTTTTTAG CAAGTTACGA ATATTTACTC 1140

AATTGAGCAG ATGTCTCTTT AAATATTTTT CTTTAATTTC TTAGCTAAGC GGAGCATCTA 1200

TCTTAAGTAT CTCTACTGAA TTTAAGACAT AATACATTTT TTTAAAAAAT CTATTAGAGT 1260

GTTTTTTCCG CACAGCGCAC ATATATCTTT TTTCTGGTAA TTCAGACAAC CTTTCTCCCG 1320

ACGATAAAAT AATATAAGAT TAACTCCTTG AACTAATTTT TTATTTTTCT TTTCTTTTTA 1380

TGTTCTTTGC AGAAAGTTTC TTATGGTCTT TTGTGAAAAG TACATTCTAT GATAATTTTT 1440

TGGCAACTCA TATAAATTTA TATATATTCC ATGTAGTTAT AAGTTAAAAA AAGCTTCCTA 1500

TTAATTCCAA GATAGAGGTT CATTTTTATA GTTTGGGCAT CCATGAGTTT TTGAAAATGT 1560

CAGAAATTTT GTTGAGTTAA TTTTACTTAC CAACTTTTAT GGCGTCATGC AGTGATCTTG 1620

GGAATGCAGC ATTATCTGGT CAATTGGTTC CTCTTGGCCA GTTGAAAAAT TTACAATACT 1680

TGTAAGACCA TATCACTTGG AATGCTTTAG TTTTTATACA GCACAATGCT TTCAATATCT 1740

GTTAAAAGTG TGAAAAAGTT GACTTTCTAG CTTCAGCAGT TGTTCGGATA ATATCTATGA 1800

AGCACTTAAA AGGCTGGGCA ATTTTTTTGT TATTATTTCA AATATTGTTA ATTGTTACTA 1860

CTTAATATGA TAAACTGATT TAACTCCTCA TGATTGGTCT CAGTCCAATG TGCCCTCATT 1920

AGTCACATNA TAAAATTGGN GGGTTGGACA AATATAACTT CTTTTCTTAA GGTCCAGAAA 1980

GAGCACTTAT CAACCTTGTC TAGCGCATAA CGTCACAGTG GGTCAGTCAC GGGCTATCCA 2040

GTTTGGGGAG GTTTTAATGA GCACTTATTT ACCTTGTCTT TTAAACGTCT GAGGATGTTA 2100

TTAAAGTCTG CATCATTCAG AGTTTAAATT AGCACTTTCA GTTGTATTAT GAATGGTACA 2160

TGAAAGATAC ATATCTTAAT GTTCCTATGC CTGTTTCAAC ATGTCTCTAA TATTCTGTTA 2220

TCTTTGTCAT CTTAAAAATG GCACTGATTA AAATGTGAGA AAGGTAGTCT TCCAATACCA 2280

TTTCATGTAT ACCAGAGAAT ATCATAATTT TTTTAAATCA TAAGTTGGGC CCTAGAGTTT 2340

TCTCAGTATT GGTCTATTTA TATTTTCCAC CATTTAGAAC TGTGTTGTCA GATGAAAATC 2400

TTGGACTTCC ACAGAAGATC TTATAGTAAA AGTATTCTTT AGATCTGATG ATGAAAGTTG 2460

TCATGGTGTG GCCTGTCCCA GAATTTAAAT CAATCCCATG TCACATGTTT GTTGATCTGA 2520

CTACTCACTG TTAATCGAAG AGTAACTATT TGTGAATTAA ATGCTTTTTT TTTTGTTCTT 2580

CATGCTTAGC GTTATAAAGG TCTACGTCTG ACTATGGTTT TTAACATGTT ATACTTTTGT 2640

ACTGACAAGT TTAAAGTTTC TCTTGTTTAC GAATTAAGAA TATATAATAT AAAACGCTTT 2700

AACTTTCTCT GTGGAAGGTG TTCTTACCTT TTTATATATA TATATAGATA CTCAGACTCT 2760

GCTGGCAATT ATATCTTACG AACTTACGAG TATACAGAAC TTGTATATTA GGTTCAGATG 2820

AGTGGCTGTA GTAGAACACC TTAAGCAAGA ACTTAATCAT GAGGTTTCAA CCTTTTAACT 2880

TTCTTTTTAG ATTTTTTCAA GTTTATGGAA AATTGTACCT CATGATCGTG GTTTCTTTCC 2940

ATAAACTTTC CATATAAGTC CGTTTCTTGA CGTTTTCATG TAAGCTGTTG ACGAGTGATT 3000

ATTAGCGGTT CTTTCAATAA TCATAATGTG TCTCACTTTG ATGAGGCCTG TACTTATTAT 3060

TGCACCTTGC ACTTAACCTT GATCCTCATG TCATCTTGAT TGTCATAGTC TACTAACCGA 3120

GTTGAACATG GTTTATCATG TCTTTTGAGG TAACAATGTA GCTTTCACCT CTGTCCTTGA 3180

TATAGGTTTA AGGCTTGCAC CTCCCACTAG CCTTTCGTTG TTTTATTCAC AGTTCACACA 3240

CCTACTAGCA CTGTTCACCT CTAGTCTTTT GTCCGCAAAT AGTAAGAAGT TTCTTTCGCA 3300

TAATAGTGGA TGATCATTTA AGAAATAGTG AATCAAATTA TCGTGTTATT GTGTTTGTAC 3360

TTTGGAATTA AATGAGTTGC TGAACATTGT TGCTGTTTAT CGTTGTCAAG GCTTTGCCAA 3420

GGAAGGCGAT TAGTAAGAGT GGGCATCCAA GCGCCTTTAT CTTGAAGGGG CGGGCGGCAC 3480

GTTGTGGATT CTGGGTGTCT ATTAGAGGAC ATTATCTATA TATACTGATT ATTTATTAGA 3540

ATATAAATCA ACTACTATAT TTTTCTTTGT AATGTTTATA TAGAAATCCC ACTCGTAAAC 3600

TTGACAAATA CCATTGAAAT ATTTGAACCT AATTAATTAG TAGTGTCAGG TTTAAATTCA 3660

AACTCATTTA ATTTTACTTT AAAAAATAAT TCTATATGAA TCGTAACAGT ATAAATATAT 3720

TAAATTACAT GTATGTGTGC CTATATATAG CTGAATGTCT AATAGACTCC AAGACGGCTG 3780

CTCTTACTGC CTAGGCGTCC AGGCAGTTCA CTGATGCTTA CCTTGACAAA TATGGGGTTC 3840

GTATGACATT GTTGGGGATC CCTATCACTG GATTCCTGTT TTGCTGACCC TCTGTTCAAT 3900

TGATTTTCAT TGATGTAGTA TTACTAGTTT TATAAATATT CTTTATTGCA ATAATTTAAC 3960

TGGAGTTTAA CAATGACAGG GAGCTTTACA GCAATAACAT AAGTGGACCA ATTCCTAGTG 4020

ATCTTGGGAA TCTGACAAAT TTGGTGAGCT TGGACCTATA CATGAATAGC TTCTCTGGAC 4080

CTATACCGGA CACATTAGGA AAGCTTACAA GGCTAAGATT CTTGTATGAC TACAAATCTT 4140

CACTAGTTTT TAACTTAATG CAATTTGATT ATCCTTTCAA GTGATTGATT ATATCACAAA 4200

TTACTGGATA GGCGTCTCAA CAACAACTGC CTCTCTGGTC CAATTCCAAT GTCACTGACT 4260

AATATTACAA CTCTTCAAGT CCTGTAAGTA TTCCGACCTT TCCAGATAGT TTTGTTGTTG 4320

TGGATGTTTC AATTTTAATA CTAAATATGT TCATCAGGGA TTTATCAAAC AATCGGCTAT 4380

CAGGACCAGT ACCGGATAAT GGCTCATTTT CTTTGTTTAC ACCTATCAGG TTTAATGCTA 4440

GTAATATCTT TAATATTATG GTTCTTACTT CTACTGCGAA AGCTATGATA ATATTTTTTT 4500

TCTCCTTCAT ATATTATCAC TTTCGCAGTT TTGGCAATAA TTTGAATTTA TGTGGACCTG 4560

TAACTGGGAG GCCCTGCCCT GGATCTCCCC CATTTTCTCC ACCACCTCCG TTCATCCCAC 4620

CATCAACAGT ACAGCCTCCA GGTGATTTAG TTTTTATATT AATTCCCGTA ATTAATTTTA 4680

TGACTGTAAA AATTGGTGTT AATTTCACCA GTTGCGAATA AAGTATTTTC CTTCTTTCTC 4740

TTCTTATTAT TATGAAGGAC AAAATGGTCC CACTGGAGCT ATTGCTGGGG GAGTAGCTGC 4800

TGGTGCTGCT TTACTGTTTG CTGCACCTGC AATGGCATTT GCATGGTGGC GGAGAAGAAA 4860

ACCGCGAGAA CATTTCTTTG ATGTGCCAGG TTAGTCCTGT AAATAGATAT CTATTGAAGC 4920

GCTTACTGTC TGTGGACTTT GTTTTCACTG TCATTAGTTA ACTTCAGCTG AAGAGGACCC 4980

AGAAGTGCAC CTTGGTCAAC TGAAGAGGTT TTCTCTGCGA GAATTGCAAG TCGCAACGGA 5040

TACTTTTAGT ACCATCCTTG GAAGAGGTGG ATTTGGTAAG GTGTATAAGG GACGCCTTGC 5100

TGATGGCTCA CTTGTAGCAG TTAAAAGGCT TAAAGAAGAA CGAACACCAG GTGGCGAGCT 5160

GCAGTTTCAA ACAGAAGTGG AAATGATTAG CATGGCTGTG CATCGAAATC TTCTGCGTCT 5220

ACGTGGTTTC TGCATGACAC CTACCGAGCG GCTTCTTGTA TATCCATACA TGGCTAATGG 5280

AAGTGTTGCG TCATGTTTAA GAGGTATCTC AGTTACAATT ACCATAACTT GCCAGAAGTT 5340

TGTTTGATTA AAAATGAAAT ATAACTCCCT ACACTATGTT AAGGTGTTAT AATTTCTGAG 5400

CAGATCTTAT TTCCCATTGC AAGATACCAG TTATTATTGT TTTTTCTGTA ATTGATACCG 5460

GTTATATTTC TTTCTTGTAT TTGGTTATAT GCAAGGATTT CGAGTCTAAT AAGTTATCAA 5520

ACTGGATGCT ATGTTTATTC TGCAATTGAA TTCTTGCTTC ATGTGCCAAA ATATATATGA 5580

TTCAACTTGG AATCATCTTA TAATATACTG TGTAAAGTCA GCTGTTGACT TTCATCATTA 5640

ATTAGTCTTC ATAAATCAGA ATCTGCCTAG TGAGCTTTAC CGACATACTC TAAACCTTTC 5700

TTATGGCCCT GTATATAATC GTCCCACTTA CTTTATTCAG TTTGTCTGCT CTCTGAATTT 5760

TTGATCTGTA CATTGTGATG TCTTGTTTTC ATCAAATGTA GAGCGTCAGC CATCAGAACC 5820

TCCCCCTGAT TGGCCAACTA GGGAGAGGAT TGCACTAGGA TCTTCTAGGG GCCTATCTAA 5880

ATTGCATGAC CATTGTGATC CCAAGATTAT CCATCGCGAT GTAAAAGCTG CAAATATATT 5940

ATTGGACGAA GAATTTGAGG CTGTTGTAGG TGATTTTGGG TTAGCTAGGC TCATGGATTA 6000

CAAGGATACC CATGTTACGA CTGCTGTAAG GGGTACCATT GGGCACATAG CTCCCGAGTA 6060

CCTCTCGACT GGAAAGTCAT CAGAGAAGAC CGATGTCTTT GGTTATGGGA TAATGCTCCT 6120

AGAGCTCATT ACTGGACAGA GGGCTTTTGA TCTTGCTCGC CTTGCGAACG ATGATGATGT 6180

TATGTTGTTG GATTGGGTAT GTGTCCCGGG TGTTCCTTTG GTTAATTATT TCACATATTA 6240

GTGCTTACTA CTTTGTTGTG GCCCTTTGTT TTTATTTCCT GCCTGTATTT GATTCTTAGT 6300

CATGTTATGC ATATTGACCT GCTTTGCAAT GTCTTTTAGG TTAAAAGCCT TTTGAAAGAG 6360

AAAAAGTTGG AGATGCTGGT CGATCCTGAC CTGCAGAACA ATTACATTGA CACAGAAGTT 6420

GAGCAGCTTA TTCAAGTAGC ATTACTCTGT ACCCAGGGTT CGCCAATGGA GCGGCCTAAG 6480

ATGTCAGAGG TAGTCCGAAT GCTTGAAGGT GATGGCCTTG CAGAAAAGTG GGACGAGTGG 6540

CAAAAAGTTG AAGTCATCCA TCAAGACGTA GAATTAGCTC CACATCGAAC TTCTGAATGG 6600

ATCCTAGACT CGACAGATAA CTTGCATGCT TTTGAATTAT CTGGTCCAAG ATAAACAGCA 6660

TATAAAATGT AATGAAATTA ATATTTTTTA TGGTT 6695

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1815 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 94..1752

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
GACAAATACC ATTGAAATAT TTGAACCTAA TTAATTAGTA GTGTCAGGTT TAAATTCAAA      60

CTCATTTAAT TTTACTTTAA AAAATAATTC TAT ATG AAT CGT AAC AGT ATA AAT      114
                                     Met Asn Arg Asn Ser Ile Asn
                                       1               5

ATA TTA AAT TAC ATG CAG TTC ACT GAT GCT TAC CTT GAC AAA TAT GGG      162
Ile Leu Asn Tyr Met Gln Phe Thr Asp Ala Tyr Leu Asp Lys Tyr Gly
        10                  15                  20

GTT CTT ATG ACA TTG GAG CTT TAC AGC AAT AAC ATA AGT GGA CCA ATT      210
Val Leu Met Thr Leu Glu Leu Tyr Ser Asn Asn Ile Ser Gly Pro Ile
    25                  30                  35

CCT AGT GAT CTT GGG AAT CTG ACA AAT TTG GTG AGC TTG GAC CTA TAC      258
Pro Ser Asp Leu Gly Asn Leu Thr Asn Leu Val Ser Leu Asp Leu Tyr
40                  45                  50                  55

ATG AAT AGC TTC TCT GGA CCT ATA CCG GAC ACA TTA GGA AAG CTT ACA      306
Met Asn Ser Phe Ser Gly Pro Ile Pro Asp Thr Leu Gly Lys Leu Thr
                60                  65                  70

AGG CTA AGA TTC TTG CGT CTC AAC AAC AAC AGC CTC TCT GGT CCA ATT      354
Arg Leu Arg Phe Leu Arg Leu Asn Asn Asn Ser Leu Ser Gly Pro Ile
            75                  80                  85

CCA ATG TCA CTG ACT AAT ATT ACA ACT CTT CAA GTC CTG GAT TTA TCA      402
Pro Met Ser Leu Thr Asn Ile Thr Thr Leu Gln Val Leu Asp Leu Ser
        90                  95                  100

AAC AAT CGG CTA TCA GGA CCA GTA CCG GAT AAT GGC TCA TTT TCT TTG      450
Asn Asn Arg Leu Ser Gly Pro Val Pro Asp Asn Gly Ser Phe Ser Leu
    105                 110                 115

TTT ACA CCT ATC AGT TTT GCC AAT AAT TTG AAT TTA TGT GGA CCC GTA      498
Phe Thr Pro Ile Ser Phe Ala Asn Asn Leu Asn Leu Cys Gly Pro Val
120                 125                 130                 135

ACT GGG AGG CCC TGC CCT GGA TCT CCC CCA TTT TCG CCA CCA CCT CCG      546
Thr Gly Arg Pro Cys Pro Gly Ser Pro Pro Phe Ser Pro Pro Pro Pro
                140                 145                 150

TTC ATC CCA CCA TCA ACA GTA CAG CCT CCA GGA CAA AAT GGT CCC ACT      594
Phe Ile Pro Pro Ser Thr Val Gln Pro Pro Gly Gln Asn Gly Pro Thr
            155                 160                 165

GGA GCT ATT GCT GGG GGA GTA GCT GCT GGT GCT GCT TTA CTG TTT GCT      642
Gly Ala Ile Ala Gly Gly Val Ala Ala Gly Ala Ala Leu Leu Phe Ala
        170                 175                 180

GCA CCT GCA ATG GCA TTT GCA TGG TGG CGG AGA AGA AAA CCG CGA GAA      690
Ala Pro Ala Met Ala Phe Ala Trp Trp Arg Arg Arg Lys Pro Arg Glu
    185                 190                 195
```

```
CAT TTC TTT GAT GTG CCA GCT GAA GAG GAC CCA GAA GTG CAC CTT GGT        738
His Phe Phe Asp Val Pro Ala Glu Glu Asp Pro Glu Val His Leu Gly
200                 205                 210                 215

CAA CTG AAG AGG TTT TCT CTG CGA GAA TTG CAA GTC GCA ACG GAT ACT        786
Gln Leu Lys Arg Phe Ser Leu Arg Glu Leu Gln Val Ala Thr Asp Thr
                220                 225                 230

TTT AGT ACC ATA CTT GGA AGA GGT GGA TTT GGT AAG GTG TAT AAG GGA        834
Phe Ser Thr Ile Leu Gly Arg Gly Gly Phe Gly Lys Val Tyr Lys Gly
            235                 240                 245

CGC CTT GCT GAT GGC TCA CTT GTA GCA GTT AAA AGG CTT AAA GAA GAA        882
Arg Leu Ala Asp Gly Ser Leu Val Ala Val Lys Arg Leu Lys Glu Glu
        250                 255                 260

CGA ACA CCA GGT GGT GAG CTG CAG TTT CAA ACA GAG GTG GAA ATG ATT        930
Arg Thr Pro Gly Gly Glu Leu Gln Phe Gln Thr Glu Val Glu Met Ile
    265                 270                 275

AGC ATG GCT GTG CAT CGA AAT CTT CTG CGT CTA CGT GGT TTC TGC ATG        978
Ser Met Ala Val His Arg Asn Leu Leu Arg Leu Arg Gly Phe Cys Met
280                 285                 290                 295

ACA CCA ACA GAG CGG CTT CTT GTA TAT CCA TAC ATG GCT AAT GGA AGT       1026
Thr Pro Thr Glu Arg Leu Leu Val Tyr Pro Tyr Met Ala Asn Gly Ser
                300                 305                 310

GTT GCG TCG TGT TTA AGA GAG CGT CAG CCA TCA GAA CCT CCC CTT GAT       1074
Val Ala Ser Cys Leu Arg Glu Arg Gln Pro Ser Glu Pro Pro Leu Asp
            315                 320                 325

TGG CCA ACT AGG AAG AGG ATT GCA CTA GGA TCT GCT AGG GGG CTT TCT       1122
Trp Pro Thr Arg Lys Arg Ile Ala Leu Gly Ser Ala Arg Gly Leu Ser
        330                 335                 340

TAT TTG CAT GAC CAT TGT GAT CCC AAG ATT ATC CAT CGT GAT GTA AAA       1170
Tyr Leu His Asp His Cys Asp Pro Lys Ile Ile His Arg Asp Val Lys
    345                 350                 355

GCT GCA AAT ATA TTA TTG GAC GAA GAA TTT GAG GCT GTT GTA GGT GAT       1218
Ala Ala Asn Ile Leu Leu Asp Glu Glu Phe Glu Ala Val Val Gly Asp
360                 365                 370                 375

TTT GGG TTA GCT AGG CTC ATG GAT TAC AAG GAT ACC CAT GTT ACA ACT       1266
Phe Gly Leu Ala Arg Leu Met Asp Tyr Lys Asp Thr His Val Thr Thr
                380                 385                 390

GCT GTA AGG GGT ACC TTG GGC TAC ATA GCT CCC GAG TAC CTC TCG ACT       1314
Ala Val Arg Gly Thr Leu Gly Tyr Ile Ala Pro Glu Tyr Leu Ser Thr
            395                 400                 405

GGA AAG TCA TCA GAG AAG ACC GAT GTC TTT GGT TAT GGG ATT ATG CTC       1362
Gly Lys Ser Ser Glu Lys Thr Asp Val Phe Gly Tyr Gly Ile Met Leu
        410                 415                 420

TTA GAG CTC ATT ACT GGA CAG AGA GCT TTT GAT CTT GCT CGC CTT GCG       1410
Leu Glu Leu Ile Thr Gly Gln Arg Ala Phe Asp Leu Ala Arg Leu Ala
    425                 430                 435

AAC GAT GAT GAT GTT ATG TTG TTG GAT TGG GTT AAA AGC CTT TTG AAA       1458
Asn Asp Asp Asp Val Met Leu Leu Asp Trp Val Lys Ser Leu Leu Lys
440                 445                 450                 455

GAG AAA AAG TTG GAG ATG CTG GTC GAT CCT GAC CTG GAG AAC AAT TAC       1506
Glu Lys Lys Leu Glu Met Leu Val Asp Pro Asp Leu Glu Asn Asn Tyr
                460                 465                 470

ATT GAC ACA GAA GTT GAG CAG CTT ATT CAA GTA GCA TTA CTC TGT ACC       1554
Ile Asp Thr Glu Val Glu Gln Leu Ile Gln Val Ala Leu Leu Cys Thr
            475                 480                 485
```

```
CAG GGT TCG CCA ATG GAG CGG CCT AAG ATG TCA GAC GTA GTC CGA ATG      1602
Gln Gly Ser Pro Met Glu Arg Pro Lys Met Ser Glu Val Val Arg Met
        490                 495                 500

CTT GAA GGT GAT GGC CTT GCA GAA AAG TGG GAC GAG TGG CAA AAA GTA      1650
Leu Glu Gly Asp Gly Leu Ala Glu Lys Trp Asp Glu Trp Gln Lys Val
    505                 510                 515

GAA GTC ATC CAT CAA GAC GTA GAA TTA GCT CCA CAT CGA ACT TCT GAA      1698
Glu Val Ile His Gln Asp Val Glu Leu Ala Pro His Arg Thr Ser Glu
520                 525                 530                 535

TGG ATC CTA GAC TCG ACA GAT AAC TTG CAT GCT TTT GAA TTA TCT GGT      1746
Trp Ile Leu Asp Ser Thr Asp Asn Leu His Ala Phe Glu Leu Ser Gly
                540                 545                 550

CCA AGA TAAACAGCAT ATAAAATGTG AATGAAATTA ATATTTTTTA TGGTTAAAAA      1802
Pro Arg

AAAAAAAAAA AAA                                                      1815
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 553 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Met Asn Arg Asn Ser Ile Asn Ile Leu Asn Tyr Met Gln Phe Thr Asp
 1               5                  10                  15

Ala Tyr Leu Asp Lys Tyr Gly Val Leu Met Thr Leu Glu Leu Tyr Ser
            20                  25                  30

Asn Asn Ile Ser Gly Pro Ile Pro Ser Asp Leu Gly Asn Leu Thr Asn
        35                  40                  45

Leu Val Ser Leu Asp Leu Tyr Met Asn Ser Phe Ser Gly Pro Ile Pro
    50                  55                  60

Asp Thr Leu Gly Lys Leu Thr Arg Leu Arg Phe Leu Arg Leu Asn Asn
65                  70                  75                  80

Asn Ser Leu Ser Gly Pro Ile Pro Met Ser Leu Thr Asn Ile Thr Thr
                85                  90                  95

Leu Gln Val Leu Asp Leu Ser Asn Asn Arg Leu Ser Gly Pro Val Pro
            100                 105                 110

Asp Asn Gly Ser Phe Ser Leu Phe Thr Pro Ile Ser Phe Ala Asn Asn
        115                 120                 125

Leu Asn Leu Cys Gly Pro Val Thr Gly Arg Pro Cys Pro Gly Ser Pro
    130                 135                 140

Pro Phe Ser Pro Pro Pro Pro Phe Ile Pro Pro Ser Thr Val Gln Pro
145                 150                 155                 160

Pro Gly Gln Asn Gly Pro Thr Gly Ala Ile Ala Gly Gly Val Ala Ala
                165                 170                 175

Gly Ala Ala Leu Leu Phe Ala Ala Pro Ala Met Ala Phe Ala Trp Trp
            180                 185                 190

Arg Arg Arg Lys Pro Arg Glu His Phe Phe Asp Val Pro Ala Glu Glu
        195                 200                 205
```

Asp Pro Glu Val His Leu Gly Gln Leu Lys Arg Phe Ser Leu Arg Glu
210 215 220

Leu Gln Val Ala Thr Asp Thr Phe Ser Thr Ile Leu Gly Arg Gly Gly
225 230 235 240

Phe Gly Lys Val Tyr Lys Gly Arg Leu Ala Asp Gly Ser Leu Val Ala
245 250 255

Val Lys Arg Leu Lys Glu Glu Arg Thr Pro Gly Gly Glu Leu Gln Phe
260 265 270

Gln Thr Glu Val Glu Met Ile Ser Met Ala Val His Arg Asn Leu Leu
275 280 285

Arg Leu Arg Gly Phe Cys Met Thr Pro Thr Glu Arg Leu Leu Val Tyr
290 295 300

Pro Tyr Met Ala Asn Gly Ser Val Ala Ser Cys Leu Arg Glu Arg Gln
305 310 315 320

Pro Ser Glu Pro Pro Leu Asp Trp Pro Thr Arg Lys Arg Ile Ala Leu
325 330 335

Gly Ser Ala Arg Gly Leu Ser Tyr Leu His Asp His Cys Asp Pro Lys
340 345 350

Ile Ile His Arg Asp Val Lys Ala Ala Asn Ile Leu Leu Asp Glu Glu
355 360 365

Phe Glu Ala Val Val Gly Asp Phe Gly Leu Ala Arg Leu Met Asp Tyr
370 375 380

Lys Asp Thr His Val Thr Thr Ala Val Arg Gly Thr Leu Gly Tyr Ile
385 390 395 400

Ala Pro Glu Tyr Leu Ser Thr Gly Lys Ser Ser Glu Lys Thr Asp Val
405 410 415

Phe Gly Tyr Gly Ile Met Leu Leu Glu Leu Ile Thr Gly Gln Arg Ala
420 425 430

Phe Asp Leu Ala Arg Leu Ala Asn Asp Asp Asp Val Met Leu Leu Asp
435 440 445

Trp Val Lys Ser Leu Leu Lys Glu Lys Lys Leu Glu Met Leu Val Asp
450 455 460

Pro Asp Leu Glu Asn Asn Tyr Ile Asp Thr Glu Val Glu Gln Leu Ile
465 470 475 480

Gln Val Ala Leu Leu Cys Thr Gln Gly Ser Pro Met Glu Arg Pro Lys
485 490 495

Met Ser Glu Val Val Arg Met Leu Glu Gly Asp Gly Leu Ala Glu Lys
500 505 510

Trp Asp Glu Trp Gln Lys Val Glu Val Ile His Gln Asp Val Glu Leu
515 520 525

Ala Pro His Arg Thr Ser Glu Trp Ile Leu Asp Ser Thr Asp Asn Leu
530 535 540

His Ala Phe Glu Leu Ser Gly Pro Arg
545 550

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTTTTTTTTT TGC 13

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGGATCTAAG 10

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ACACGTGGTC 10

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCAGCACAGG 10

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTTTTTTTTT TCTG 14

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTTTTTTTTT TCA 13

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GACATCGTCC 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCCTACTGGT 10

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ACACGTGGTC 10

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGTGACTGTC 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TCTTGGACCA GATAATTC 18

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CTCTGATGAC TTTCCAGTC 19

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AATGGCATTT GCATGG 16

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Ser Pro Pro Pro Pro
1 5

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

His Arg Asp Val Lys Ala Ala Asn
1 5

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Daucus carota

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Gly Thr Leu Gly Tyr Ile Ala Pro Glu
1 5

REFERENCES

Aleith, F. and Richter, G. (1990) Planta 183, 17-24.

Bent, A.F., Kunkel, B.N., Dahlbeck D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B.J. (1994) Science 265, 1856-1860.

Braun T, Schofield P.R. and Sprengel R. (1991) EMBO J. 10, 1885-1890.

Dangl, J.L. (1995) Cell 80, 363-366.

De Jong, A.J., Schmidt, E.D.L. and De Vries, S.C. (1993) Plant. Mol. Biol. 5, 367-377.

De Vries, S.C., Booij, H., Meyerink, P., Huisman, G., Wilde, H.D., Thomas T.L. and Van Kammen, A. (1988a) Planta 176, 196-204.

De Vries, S.C., Hoge, H., and Bisseling, T. (1988b) In Plant Molecular Biology Manual, B6, S.B. Gelvin, R.A. Schilperoort, and D.P.S. Verma. eds (Dordrecht, the Netherlands: Kluwer Academic Publishers), pp 1-13.

Dubois, T., Guedira, M., Dubois, D. and Vasseur, J. (1991) Protoplasma 162, 120-127.

Dudits, D., Györgyey, J., Bögre, L. and Bakó, L. (1995) In: In vitro Embryogenesis in Plants. Ed. Thorpe, T.A. Kluwer Press, pp. 267-308.

Engler J.A., Van Montagu M. and Engler G. (1994) Plant Mol. Biol. Rep. 12, 321-331.

Gamborg, O.L., Miller, R.A. and Ojima, K. (1968) Exp. Cell. Res. 50, 151-158.

Giorgetti, L., Vergara, M.R., Evangelista, M., LoSchiavo, F., Terzi, M. and Ronchi, V.N. (1995) Mol. Gen. Genet. 246, 657-662.

Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989) Cell 56, 149-160.

Goldberg, R.B., Barker, S.J., Perez-Grau, L. (1989) Cell 56, 149-160.

Goldberg, R.B., de Paiva, G. and Yadegari, R. (1994) Science 266, 605-614.

Govind, S. and Steward, R. (1991) Dorsoventral pattern formation in *Drosophila.* Trends Genet. 7, 119-125.

Guzzo, F., Baldan, B., Levi, M., Sparvoli, E., LoSchiavo, F., Terzi, M., and P. Mariani (1995). Protoplasma 185, 28-36.

Guzzo, F., Baldan, B., Mariani, P., LoSchiavo, F. and Terzi, M. (1994) Exper. Botany 45, 1427-1432.

Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42-52.

Hashimoto, C., Hudson, K.L. and Anderson K.V. (1988) Cell 52, 269-279.

Heck, G.R., Perry, S.E., Nichols, K.W. and Fernandez, D.E. (1995) AGL15, a MADS domain protein expressed in developing embryos. Plant Cell 7, 1271-1282.

Heldin, C-H. (1995) Cell 80, 213-223.

Hodge, R., Paul, Wyatt, Draper, J. and Scott, R. (1992) Plant J. 2, 257-260.
Horn, M.A. and Walker, J.C. (1994) Biochim. Biophys. Acta 1208, 65-74.
Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J. and Jones, J.D.G. (1994) Science 266, 789-792.
Kobe B. and Deisenhofer J. (1994) TIBS 19, 415-421.
Li, F., Barnathan, E.S. and Karikó, K. (1994) Nucl. Acid Res. 22, 1764-1765.
Liang, P. and Pardee, A.B. (1992) Science, 257, 967-971.
Meyerowitz, E.M. (1995) EDBC 95 congress of the European developmental biology organization. Abstract SI4.
Mu J-H., Lee H-S. and Kao T-h. (1994) Plant Cell 6, 709-721.
Pennell, R.I., Janniche, L., Scofield, G.N., Booij, H., de Vries, S.C. and Roberts, K. (1992) J. Cell Biol. 119, 1371-1380.
Rounsley, S.D., Ditta, G.S. and Yanofsky, M.F. (1995) Plant Cell 7, 1259-1269.
Sato, S., Toya, T., Kawahara, R., Whittier, R.F., Fukuda, H. and Komamine, A. (1995) Plant Mol. Biol. 28, 39-46.
Shelton, C.A. and Wasserman, S.A. (1993) Cell 72, 515-525.
Sterk, P., and De Vries, S.C. (1992) In Redenbaugh K (ed), Synseeds: Applications of synthetic seeds to crop improvement, CRC Press, London (1992).
Sterk, P., Booij, H., Schellekens, G.A., Van Kammen, A. and De Vries, S.C. (1991) Plant Cell 3, 907-921.
Thomas, T.L. (1993) Plant Cell 5, 1401-1410
Toonen, M.A.J. and De Vries, S.C. (1995) In: Embryogenesis, the generation of a plant. Eds. Wang, T.L. and Cuming, A. BIOS Scientific Publishers, Oxford, UK, pp. 173-189.
Toonen, M.A.J., Hendriks, T., Schmidt, E.D.L., Verhoeven, H.A., Van Kammen, A and De Vries, S.C. (1994) Planta 194, 565-572.
Toonen, M.A.J., Schmidt, E.D.L., Hendriks, T., Verhoeven, H.A., Van Kammen, A. and De Vries, S.C. (1996) *submitted.*
Torii, K.U. and Komeda, Y. (1994) 4th international congress of plant molecular biology. Abstract 692,
Van Engelen, F.A. and De Vries, S.G. (1992) Trends Genet. 8, 66-70.
Varner, J.E. and Lin, L.-S. (1989) Cell 56, 231-239.
Walker J.C. (1994) Plant Mol. Biol. 26, 1599-1609.
Walker, J.C. (1993) Plant J. 3, 451-456.
Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C. and Baker, B. (1994) Cell 78, 1101-1115.
Wilde, H.T., Nelson, W.S., Booij H., de Vries, S.C. and Thomas, T.L. (1988) Planta 176, 205-213.
Wurtele, E.S., Wang, H., Durgerian, S., Nikolau, B.J. and Ulrich, T.H. (1993) Plant Physiol. 102, 303-312.
Zhao, Y., Feng, X.H., Watson, J.C., Bottino, P.J. and Kung, S.D. (1994) Plant Molec. Biol. 26, 791-803.
Zimmerman, J.L. (1993) Plant Cell 5, 1411-1423.

CLAIMS

1. A method of producing apomictic seeds comprising the steps of:
(i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in an active form in a cell, or membrane thereof, renders said cell embryogenic,
(ii) regenerating the thus transformed material into plants, or carpel-containing parts thereof, and
(iii) expressing the sequence in the vicinity of the embryo sac.

2. A method according to the preceding claim, wherein the apomictic seeds are of the adventitious embryony type.

3. A method according to either of the preceding claims, wherein expression of the sequence yields a protein kinase capable of spanning a plant cell membrane.

4. A method according to the preceding claim wherein the kinase is capable of autophosphorylation.

5. A method according to any of the preceding claims, wherein the protein is a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain.

6. A method according to the preceding claim, wherein the protein lacks a functional ligand binding domain but comprises a proline box, a transmembrane domain, a kinase domain and a protein binding domain.

7. A method according to any preceding claim, wherein once incorporated into the cell membrane, the protein binding domain is located intra-cellularly.

8. A method according to any preceding claim, wherein the sequence further encodes a cell membrane targeting sequence.

9. A method according to any preceding claim, wherein the sequence is that depicted in

SEQ ID Nos. 1 or 2 or is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

10. A method according to any preceding claim, wherein the sequence is modified in that known mRNA instability motifs or polyadenylation signals are removed and/or codons which are preferred by the plant into which the sequence is to be inserted are used so that expression of the thus modified sequence in the said plant yields substantially similar protein to that obtained by expression of the unmodified sequence in the organism in which the protein is endogenous.

11. A method according to any preceding claim, wherein expression of the sequence is under control of an inducible or developmentally regulated promoter.

12. A method according to the preceding claim, wherein expression of the sequence is under control of one of the following: a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*.

13. A method according to any of the preceding claims, wherein the sequence is expressed in the somatic cells of the embryo sac, ovary wall, nucellus, or integuments.

14. A method according to any of the preceding claims, wherein the endosperm within the apomictic seed results from fusion of polar nuclei within the embryo sac with a pollen-derived male gamete nucleus.

15. A method according to the preceding claim, wherein the sequence encoding the protein is expressed prior to fusion of the polar nuclei with the male gamete nucleus.

16. Recombinant DNA comprising a sequence encoding a protein having the sequence depicted in SEQ ID No. 3, or a protein substantially similar thereto which is capable of being membrane bound and which has kinase activity.

17. DNA according to the preceding claim, which further encodes a cell membrane targeting sequence.

18. DNA according to either of claims 16 or 17, wherein the protein is a leucine rich repeat receptor like kinase and comprises a ligand binding domain, a proline box, a transmembrane domain, a kinase domain and a protein binding domain, the ligand binding domain optionally being absent or functionally inactive.

19. DNA according to any one of claims 16 to 18, in which the protein encoding region is under expression control of a developmentally regulated or inducible promoter.

20. DNA according to the preceding claim, wherein the promoter is one of the following: a promoter which regulates expression of SERK genes *in planta*, the carrot chitinase DcEP3-1 gene promoter, the *Arabidopsis* AtChitIV gene promoter, the *Arabidopsis* LTP-1 gene promoter, the *Arabidopsis* bel-1 gene promoter, the petunia fbp-7 gene promoter, the *Arabidopsis* ANT gene promoter, the promoter of the O126 gene from *Phalaenopsis*.

21. DNA according to any preceding claim, having the sequence depicted in SEQ ID Nos. 1 or 2 or a sequence which is complementary to one which hybridizes under stringent conditions with the said sequences and which encodes a membrane bound protein having kinase activity.

22. DNA according to any one of claims 16 to 21, wherein the sequence is modified in that known mRNA instability motifs or polyadenylation signals are removed and/or codons which are preferred by the plant into which the DNA is to be inserted are used so that expression of the thus modified DNA in the said plant yields substantially similar protein to that obtained by expression of the unmodified DNA in the organism in which the protein is endogenous.

23. DNA which is complementary to that which hybridizes under stringent conditions with the DNA of any one of claims 16 to 22.

24. A vector containing a DNA sequence as claimed in any one of claims 16 to 23.

25. Plants transformed with the recombinant DNA of any one of claims 16 to 23 or the vector of claim 24, the progeny of such plants which contain the DNA stably incorporated, and/or the apomictic seeds of such plants or such progeny.

26. Plants transformed with the DNA comprised by the recombinant DNA of any one of claims 16 to 23.

27. Use of the DNA of any one of claims 16-23 in the manufacture of apomictic seeds.

28. Plants which result from the apomictic seeds which result from the method of any one of claims 1-15 or 26 27.

29. A method of obtaining cultivars comprising the steps of fertilizing plants with the pollen of the plants of either of claims 25, 26 or 28 and cultivars which result from the said method.

30. A method of obtaining embryogenic cells in plant material, comprising transforming the material with a recombinant DNA sequence as claimed in any one of claims 16-23, the DNA comprised by the recombinant DNA of any one of claims 16 to 23, or the vector of claim 24, expressing the sequence in the material or derivatives thereof and subjecting the said material or derivatives to a compound which acts as a ligand for the gene product of the said sequence.

31. A method according to the preceding claim, wherein the sequence encodes a leucine rich repeat receptor like kinase, and the compound is a phyto-hormone.

ABSTRACT

The present invention provides, *inter alia*, a method of producing apomictic seeds comprising the steps of:

(i) transforming plant material with a nucleotide sequence encoding a protein the presence of which in a cell, or membrane thereof, renders said cell embryogenic,

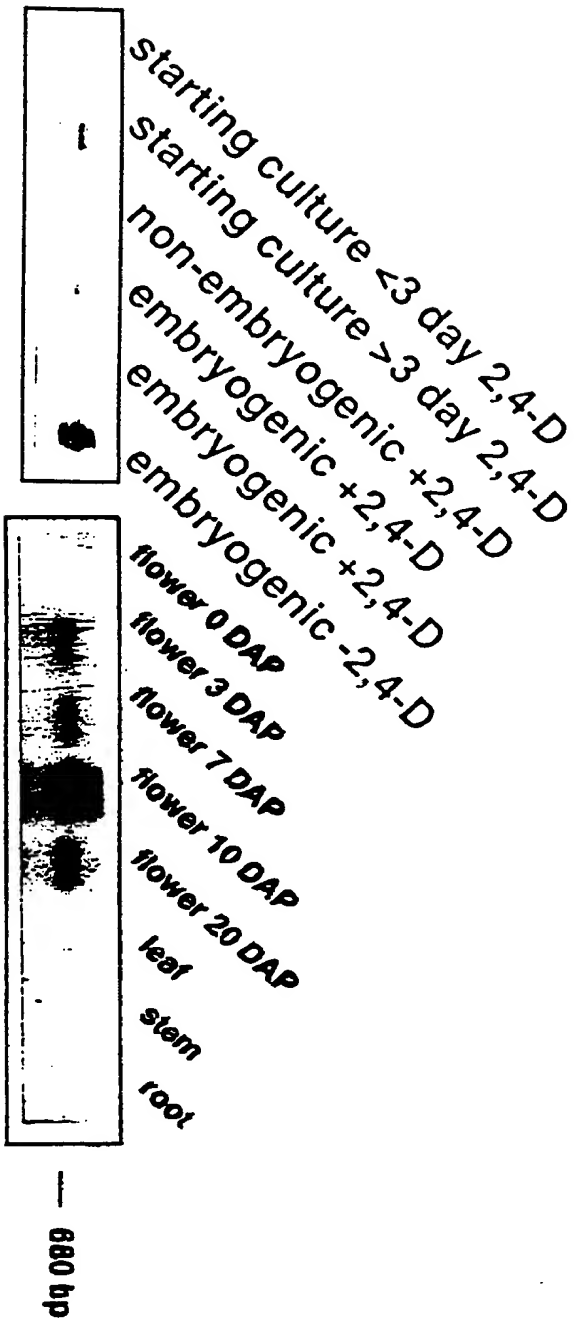
(ii) regenerating the thus transformed material into plants, or carpel containing parts thereof, and

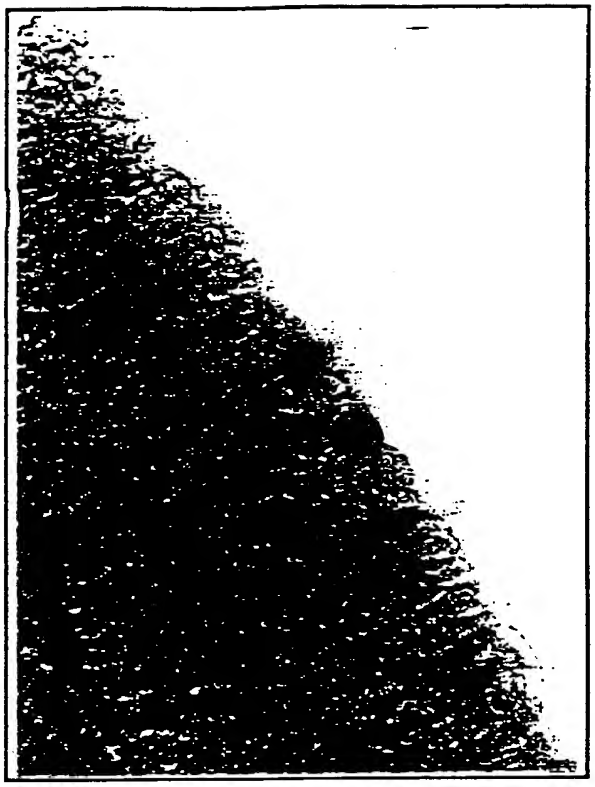
(iii) expressing the sequence in the vicinity of the embryo sac.

The protein may be a leucine repeat rich receptor kinase which preferably is modified to the extent that the ligand binding domain is deleted or functionally inactivated.

B7-1103/SG

Figure 1





A

B

Figure 2

137-1103/SG

Figure 3

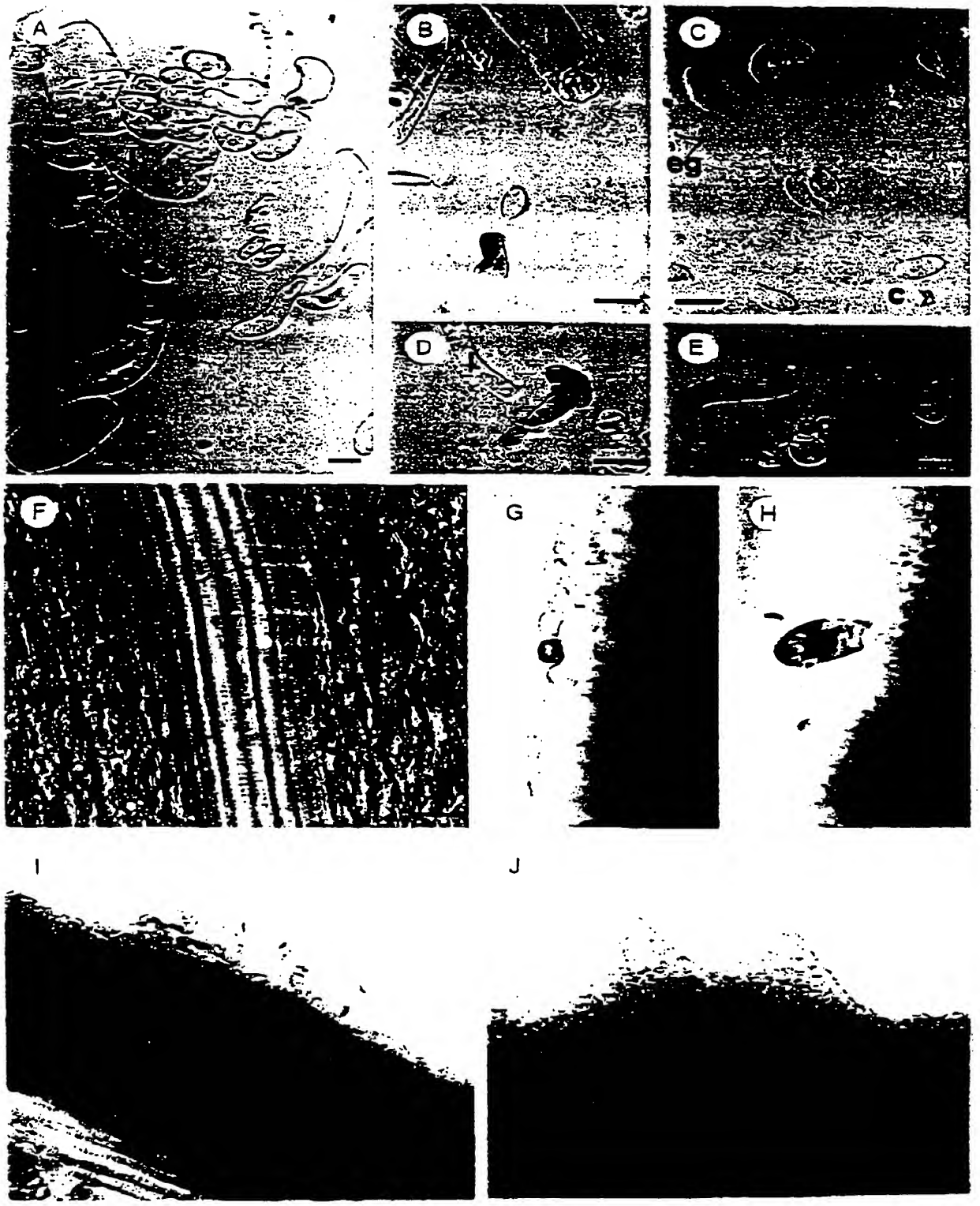


FIG 6